Interoperable, Personalized, On-demand Geospatial Data Access and Services Based on OGC Web Coverage Service (OWS) Specification

Liping Di, Wenli Yang, and Donna Deng
Laboratory for Advanced Information Technologies and Standards (LAITS)
George Mason University
9801 Greenbelt Road, Suite 316-317
Lanham, MD 20706-6200

Ken McDonald
NASA Goddard Space Flight Center, Code 423.0
Greenbelt, MD 20771

Abstract- Geospatial data is one of important sources of information for Earth system science research, environmental and natural resource management, decision-making, and other applications. However, current one-size-fit-all approach at data centers for geospatial data access and services cannot meet data users' requirements and expectations. Data users are calling for the interoperable, personalized, on-demand data access and services (IPODAS) that enables them to obtain data from multiple sources in the form exactly matching their requirements. This paper describes IPODAS and the way for providing IPODAS through OGC Web coverage service specification. It also describes IPODAS of NASA HDF-EOS data with two components of NASA HDF-EOS Web GIS Software Suite (NWGISS): the multiple protocol geoinformation client and the web coverage server.

I. INTRODUCTION

Earth System Science (ESS) research is usually multi-disciplinary in nature. It requires data collected by in-situ instruments and airborne or spaceborne remote sensors as well as the socioeconomic data. Those data are typically archived at and distributed by many different data centers. The goal of ESS research is to discover new knowledge about the Earth system. Normally, the processes of knowledge discovery in ESS involve three consecutive steps in the data and information flow*: 1) Geoquery; 2) Geodata and information assembly; and 3) Geocomputation.* Geoquery is to locate and order data from the data repositories. The geocomputation is the analysis and simulation of the complex Earth system using the data and information from the geoquery. Geodata and information assembly assembles the data and information from data centers based on the needs of geocomputation.

The geoquery is the first step in the processes of ESS knowledge discovery. Currently most of data centers provide on-line data catalogs for data users to locate the data they need. Once users find the required data, users can place an order for the data. Then data centers handle the order by pulling out data from the archive and copying the data to computer media for delivery or staging the data in a computer for user to ftp. This process normally takes days and even weeks. Some data centers have placed data on-line so that data users can directly download data instead of ordering. In such a case, users can obtain the data in real time once they find the requested data. The data obtained from data centers normally are in the pre-made archival form. Services are normally not provided for converting the user-requested data from the archival form to user-needed form. The data centers assume that data in the archival form can meet most users' requirement. For example, NASA distributes data in pre-made data granules. The data granules may cover a geographic region larger than the research area of the user. The data may be in a projection not wanted by the user. Because of the multidisciplinary nature of Earth System Science, datasets from multiple data centers are very diverse, and in many cases, the temporal and spatial coverage, resolution, format, and projections are incompatible. Therefore, after obtaining the data, scientists spend considerable time assembling the data and information into a form ready for analysis in the geocomputation step, even when the analysis is very simple [1]. Because each user has different requirements, current one-size-fit-all approach for providing data access and services to users cannot meet users' requirements. Although data and information systems are the core for supporting ESS research, current systems, which mostly

facilitate only data search and ordering, fall far short of supporting multidisciplinary ESS studies.

II. Interoperable Personalized On-demand Data Access and Services (IPODAS)

Data services processes the data on archive to produce data products on users' behavior. The most common data service functions include spatial, temporal, and parameter-based subsetting and subsampling, georectification, reprojection, and reformatting. Those functions can be combined to produce data products that exactly match users requirements on data.

It is most likely in the near future there are many data servers on the web that provide data to users. In order for clients to easily access data from different servers for integration and analysis, those data and servers have to be interoperable. Therefore, it is required to have standards for the interface between the data servers and clients and the encoding for the data. Because of the diversities in research topics of ESS and in the temporal and spatial scales of the research, individual research project has its unique requirements to the data. Therefore, data requirements for each case will be different. A provider has to personalize the archival data in order to meet individual user's requirement. In order to personalize the data, data service functions have to be executed so that users can obtain data that exactly meet their requirements. Because users have their unique requirements and the data providers cannot know accurately in advance who will use the data and what are their requirements, therefore, all services have to be conducted on-demand at real time basis. If data providers can provide data access and services described in above, data users can obtain the data at real or near real time in the form that exactly matches their requirements in term of the spatial and temporal coverage, projection, data format, spatial resolution.

III. OGC Web Coverage Services (WCS)

The Open GIS Consortium, Inc. (OGC) is a not-for-profit international membership-based organization founded in 1994 to address the lack of interoperability among systems that process georeferenced data, and between them [2]. Since 1999 OGC has successfully implemented three web-based geospatial interoperability programs, including Web Mapping Testbed I (WMT-1) in 1999, WMT II in 2000, and OGC Web Service Initiative-I (OWS-1) in 2001, and produced a set of web-based data interoperability specifications [3]. One of the most important interoperability specifications for data access from OGC is the Web Coverage Service (WCS) Specification. It allows a WCS client to access real multi-dimensional, multi-temporal data from coverage servers. WCS provides an interoperable way of accessing geospatial data, especially those from remote sensing. OWS-1 developed WCS specification and produced two draft versions of WCS specification, v0.5 [4] and v0.6 [5]. Currently a new version of WCS, version 0.7 is out for comments.

IV. NASA EOSDIS Data Environment

NASA Earth Science Enterprise (ESE) is generating huge volume of remote sensing data for supporting Earth system science and application research. Most of the data are in HDF-EOS, the standard format for NASA's Earth Observing System (EOS) Data and Information System (EOSDIS). The data are archived at and distributed by Distributed Active Archive Centers (DAACs). There are two typical groups of data users in EOSDIS. The first group of users is the small numbers of well-funded scientists who have significant resources that enable them to handle large volumes of HDF-EOS data. The second group of data users is the large numbers of small data users with limited resources. For this group of data users, they normally go to NASA DAACs to search and find, and then order the data. The data center takes the order and delivers the data in computer media as data granules to users. They may also obtain data from other data providers. Once the data are in house, the users will need to do a lot of preprocessing steps before a geographic information system (GIS) can analyze the data. The preprocessing steps, including subsetting, subsampling, reprojection, reformatting, geometric rectification, etc., will take significant amount of time and resources from the data users. Because data users have different requirements for data, such as spatial and temporal coverages, the data parameters, and formats, it is not feasible to preprocess the data in advanced. The data services have to be conducted on the personalized, on-demand, real or near real time basis. Therefore, providing IPODAS to users by NASA DAACs will greatly enhance the interoperability and public use of EOS data.

V. IPODAS THROUGH NWGISS BASED ON OGC WCS SPECIFICATION

NASA HDF-EOS Web GIS Software Suite (NWGISS) is a multiple OGC protocol compliant, component-based software system that makes NASA HDF-EOS data available to broad GIS user community through the Web. The major goals for developing NWGISS are to test the implement-ability of OGC Web interoperability specifications in NASA environment, makes HDF-EOS data available to GIS applications according to OGC Web interoperability specifications, test the interoperability of geospatial data services provided by different data providers, and prototype the concept of providing interoperable personalized on-demand data access and services to users by using NWGISS and other OGC-standard compliant servers.

Currently, NWGISS consists of following components: a map server, a coverage server, a catalog server, a multiple protocol geoinformation client (MPGC), and a toolbox. All NWGISS components can work both independently and collaboratively. The OGC specifications implemented in NWGISS include the Web Map Services (WMS), Web Coverage Services (WCS), and Catalog Inter-operability Specification (CIS).

The key components of NWGISS for providing interoperable, personalized, on-demand data access and services are the NWGISS Multiple Protocol Geoinformation Client (MPGC) and the OGC WCS-compliant coverage servers, including NWGISS Web Coverage Server. The interactions between NWGISS MPGC and OGC compliant Web Coverage Servers provide IPODAS to users.

- *OGC WCS-compliant Web coverage servers*: OGC Web Coverage Services (WCS) is designed for enabling GIS clients to access multi-dimensional, multi-temporal geospatial data. The version 0.6 and 0.7 WCS specifications define three interface protocols, namely getCapabilities, getCoverage, and describeCoverageType. The specifications require the compliant servers to provide subsetting and resampling capabilities. The specifications also provide a way for a WCS server to serve the same data in multiple map projections and multiple coverage-encoding formats that client can choose. For servers providing those choices, normally they are implemented as reprojection/georectification and reformatting services provided on-the-fly based on client's request. The NWGISS Web Coverage Server is one of a few fully implemented WCS servers in the world. Currently we have implemented both version 0.5 and 0.6 of the draft WCS specification and are implementing version 0.7 specification. Three coverage-encoding formats are available in NWGISS, namely, HDF-EOS[8], GeoTIFF[9], and NITFF[10]. We also have implemented georectification for HDF-EOS swath data so that the swath data can be served as georectified (in Lat/Lon coordinate system) grid data. The detailed description of the NWGISS Web Coverage Server can be found in [11].
- *NWGISS multiple protocol geoinformation client(MPGC)*: The OGC WCS-compliant servers alone cannot provide IPODAS to users. This is because an ESE research project normally requires data from multiple servers and not all servers can provide data in user-requested projection and format, although users can retrieve those data at the same spatial and temporal coverages. Therefore, any WCS client intending to provide IPODAS has to have built-in functions for reprojection and reformatting. The NWGISS MPGC is a comprehensive multiple OGC protocol compliant client. Currently it complies with OGC web map service specification 1.1.0 and lowers and OGC web coverage service specification 0.6 and lowers. We are implementing OGC catalog interface and web feature service specifications. The NWGISS MPGC was developed from the original web coverage client. The coverage part of MPGC can interactively communicate with all OGC-compliant coverage servers (not only with NWGISS) for accessing multi-dimensional geospatial data and handling HDF-EOS, GeoTIFF and NITF coverage-encoding formats. Besides performing basic WCS client-server communication, coverage access, visualization, and user interaction, the client also provides following data service functions:
  - Georectification on un-rectified swath data
  - Reprojection/resampling for georectified grid data.
  - Subsetting and subsampling

- Reformatting to major GIS formats

The execution of those functions is automatically arranged based on the user's data requirement and the information about the data and services in servers. The interaction between NWGISS MPGC and OGC compliant web coverage servers provides interoperable, personalized, on-demand data access and services of remote sensing data.

In NWGISS MPGC, there is a user interface called project. Within it, users can build a project by providing requirements for the map projection, resolution, formats, spatial and temporal coverage, as well as the data coverage layers. When the users request the MPGC to build the project, it means MPGC will request data from WCS servers, and if necessary, perform data services functions transparently to the users to produce the data products that exactly meet users requirements. First, MPGC will check with the capabilities xml documents from servers to find if the requirements on the data can be met by the server-side functions. If there are mismatches between the data that servers can provide and the user requirements, the client will request the data from the servers in the forms they can provide, and transparently invoke data service functions in the client to convert the data in server provided forms to the users' requested form. At the end, users obtained the data in the form they want.

VI. Conclusions

The interoperable, personalized on-demand data access and services are one of important requirements of the next generation data information system for supporting Earth Science Research. The OGC WCS specification provides the standard for implementing IPODAS. The successful demonstrations and uses of NWGISS system show it is feasible for the data interoperability based on OGC WCS. It also demonstrates that providing IPODAS to data users greatly enhances the use of remote sensing data in scientific research and applications.

Acknowledgement

This project was supported mainly by grants from NASA Earth Science Technology Office (ESTO). Additional funding was provided by Open GIS Consortium (OGC) as part of OGC WMT II and OWS-I.

References

[1] L. Di, and K. McDonald, "Next Generation Data and Information Systems for Earth Sciences Research", in *Proceedings of the First International Symposium on Digital Earth*, Volumn I. Science Press, Beijing, China. p 92-101, 1999.

[2] OGC Homepage, http://www.opengis.org

[3] OGC Interoperability Program, http://www.opengis.org/ogcInterop.htm

[4] J. Evans, eds, OGC web coverage service version 0.5, http://ip.opengis.org/ows1/docs/01-018r1.doc

[5] S. Fellah, eds, OGC web coverage service schema version 0.3, http://ip.opengis.org/ows1/docs/WCS0.3.zip

[6] L. Di, W. Yang, M. Deng,, D. Deng, and K. McDonald, "The prototype NASA HDF-EOS Web GIS Software Suite (NWGISS)," *Proceedings of the NASA Earth Science Technologies Conference*. Greenbelt, Maryland. August 28-30, 2001 (CD-ROM, 4pp).

[7] J. de La Beaujardière, eds, Web Map Service Implementation Specification, Version 1.1.0, http://www.opengis.org/techno/specs/01-047r2.pdf

[8] HDF-EOS4(HE4), http://hdfeos.gsfc.nasa.gov/hdfeos/he4.cfm

[9] N. Ritter and M Ruth, GeoTIFF Format Specification GeoTIFF Revision 1.0, http://remotesensing.org/geotiff/spec/geotiffhome.html

[10] National Imagery Transmission Format Version 2.1, Natioal Imagery and Mapping Agency (NIMA), Bethesda, Maryland.

[11] L. Di and W. Yang 2002. "Serving NASA HDF-EOS Data through NWGISS Coverage Server." Proceedings of *Proceedings of the NASA Earth Science Technologies Conference*. Pasadena, CA. June 11-13. (CD-ROM).